**Course: Advanced Bioinformatics**

**Module title: Introduction to Phylogenetics**

**Module no. : 43**

Phylogenetics is defined as evolutionary relationships among groups of organisms (e.g. species, populations), which are discovered through molecular sequencing data and morphological data matrices.

Now, I will provide details of some of the required terms for understanding phylogentic analysis.

Phyle: tribe, clan, race

Genesis: origin, source, birth.

Phylogenesis: process

Phylogeny: science on this process

Phylogenetics: is phylogeny based on analysis of sequences of biological macro-molecules(DNA, RNA PS)

Phylogeny Algorithms Diversity: Different individuals carry different variants of the same basic blue print

Mutations: The DNA sequence can be changed due to single base changes, deletion, insertion of DNA segments, etc.

Motivation:

* To understand lineage of various species (evolutionary history)
* To understand how various functions evolve
* To inform multiple alignments
* To map virus strains (vaccine construction)
* To identify what is most conserved / important in some class of sequences

Historical Note:

Until mid 1950's phylogenies were constructed by experts based on their opinion (subjective criteria)

Since then, focus on objective criteria for constructing phylogenetic trees

Thousands of articles in the last decades

Important for many aspects of biology

* + Classification
  + Understanding biological mechanisms

Overview of phylogenetics

Phylogenetic analysis gives insight into how a family of related sequences has been derived during evolution. The evolutionary relationships among the sequences are shown as branches of a tree. The length and nesting of these branches reflects the degree of similarity between any two given sequences. The objective of phylogenetic analysis is to determine the length of the branches and to figure out how the tree should be drawn.

Sequences that are the most closely related are drawn as neighboring branches on a tree.

Phylogenetic analysis is dependent upon good multiple sequence alignment programs. Given a multiple sequence alignment, phylogenetic analysis tries to group sequences with similar patterns of substitutions in order to reconstruct a phylogenetic tree. For instance, consider that we have two sequences that are related. Given these two sequences, an ancestoral sequence can be (partially) derived. With more similar sequences, more information can be gathered to add to a correct derivation and evolutionary history.

Uses of phylogenetic Analysis

Given a set of genes (such as a family of genes) phylogenetic analysis can help determine which genes are likely to have equivalent functions.

Used to follow changes occurring in a rapidly changing species such as a virus. Take for instance influenza. By studying the rapidly changing genes through phylogenetic analysis, the next year’s strain can be predicted, and a flu vaccination can be developed. The prediction is not always correct, but it gives a level of protection.